DOCUMENT RESUME

ED 059 972 SP 005 571

AUTHOR Quereshi, M. Y.; Widlak, Frederick W.
TITLE Teacher Evaluation as a Function of the Students' Sex and Achievement Level.
INSTITUTION Marquette Univ., Milwaukee, Wis.
PUB DATE [71]
NOTE 12p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Academic Achievement; *Sex Differences; *Statistical Analysis; *Student Teacher Relationship; *Teacher Rating
IDENTIFIERS MARS; Michill Adjective Rating Scale; Rating Scale for Teachers; RST

ABSTRACT

This study was designed to determine how the ratings of a college teacher were influenced by the sex and achievement level of the students. Sixty-seven male and 73 female undergraduates were used in the main analyses, with an additional 72 subjects in subsidiary analyses. During the first week of the course subjects were asked to rate themselves, father, mother, and the instructor in the course on the Michill Adjective Rating Scale (MARS) which measures four relatively independent factors (unhappiness, extraversion, self-assertiveness, and productive persistence). They also rated themselves and the instructor on the same scale near the end of the course, rated the instructor on the Rating Scale for Teachers (RST) consisting of 17 itmes dealing with aspects of a teacher's job and personality, and took five teacher-made achievement tests during the semester. Students were separated into high and low achievement levels, using the median of the cutoff. Two-way multivariate analysis showed significant differences in teacher ratings on MARS and RST factors between the sexes and between high and low achievement level subjects. Also, significant sex and achievement level interaction was found for the RST data. (MBM)

ED 059972

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEH REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

Teacher Evaluation as a Function of the Students' Sex and Achievement Level

M. Y. Quereshi and Frederic W. Widlak

Marquette University Purdue University

What characteristics distinguish effective from ineffective teachers has been the subject of considerable research as evidenced by the reviews by Getzels & Jackson (1963), McKeachie (1963), and Remmers (1963). Many of these studies have been concerned with correlations between ratings and personality characteristics of the teachers, but a few (Kerlinger, 1966; Yonge & Sassenrath, 1969) have investigated teacher evaluations in relation to the personality characteristics of the raters. The latter area, although equally important, has not received its due share of attention. However, there is some indication that the effectiveness of a teacher as perceived by the students not only depends on the personality and teaching style of the teacher, but also on the personal and biosocial traits of the students themselves. For example, male and female students do not value certain personality characteristics of significant others (parents, teachers, spouses, etc.) in the same way (Kohn & Fiedler, 1961). There is also a general belief among many educators that students who do well in school view teachers differently from those who do poorly in academic work. The present study was designed to determine how the ratings of a college teacher were influenced by the sex and achievement level of the students.

Method

Subjects

Two hundred and twelve undergraduates completed a one-semester upper division course in psychological testing at Marquette University between 1966

SP005571

and 1968. However, complete data involving at least two separate testings on the instruments employed were available only on 140 Ss, 67 males and 73 females. Thus the main analyses were based on 140 Ss only, but a number of subsidiary analyses utilized the data on the total sample of 212. At the time of first testing, the mean ages for 67 males and 73 females were 263.3 and 260.0 months, with standard deviations of 21.3 and 15.0 months, respectively.

Instruments and Procedure

During the first week of the course Ss were asked to rate themselves and three significant others (father, mother, and the instructor in the course) on the Michill Adjective Rating Scale (MARS) which comprises 48 adjectives such as "ambitious", "determined", "nervous", etc. accompanied by five-point ratings ranging from 1 ("Very Untypical") to 5 ("Very Typical"). As described previously (Quereshi, 1970), the MARS measures four relatively independent factors (unhappiness, extraversion, self-assertiveness, and productive persistence) with reliability indices (based on 441 Ss) ranging between .78 and .90 and cross-sample congruence coefficients ranging between .80 and .95. Each S was provided with four copies of MARS and was asked to rate in order himself, his father, his mother, and the instructor in the course. Since Ss generally had had very limited contact with the instructor (not extending beyond two class meetings) at the time of first testing, they were asked to give their best judgments by either relying on their own "intuition" or by utilizing any information that they had received about the instructor from their friends or acquaintances who had taken the same course in the past. Ss were told not to write their names on any of the protocols, but the cover-sheet asked for such information as sex, date of birth, year in college, and father's

occupation. Ss were assured that the information would be used only for research purposes and only after Ss had graduated from college.

The second testing with MARS (requiring only rating themselves and the instructor) usually took place near the end of the course, after Ss had had the opportunity to interact with the instructor, both inside and outside the classroom, for several months, and had taken at least four achievement tests based on the lecture material and reading assignments for the course. In addition, Ss were administered the Rating Scale for Teachers (RST) consisting of 17 10-point items dealing with such aspects of a teacher's job and personality as preparation for class meetings, interest and enthusiasm in his subject, ability to arouse interest in students, organization of the course, scholarship, ability to express thought, enunciation, thinking demanded of students, assignments, leading discussion and questions, sense of proportion, feeling between instructor and students, sense of humor, self-confidence tolerance and liberality, personal appearance, and personal peculiarities. Finally, Ss were given a final examination in the course in accordance with the university regulations and schedule. The final, the four previous classroom achievement tests, and three take-home exercises were employed to determine the final grade in the course as well as the achievement level for the purpose of the present study. Except for the MARS, all other data were collected as part of a regular course and constituted an in situ experiment. The course was essentially of the lecture type conducted by a male faculty member of the psychology department, about 40 years old, with a Ph. D. degree and post-doctoral teaching and research experience of about 10 years at the conclusion of this study. In conducting the course, the teacher emphasized the acquisition, critical evaluation, and application of facts and principles and provided specific guidelines about course requirements and expectations at the

beginning of the course prior to the administration of MARS.

The classroom achievement tests were constructed by the instructor by sampling items from published and unpublished sources as well as by writing new items. The correlation between the sum of scores on classroom tests 1 and 2 and that of tests 3 and 4, based on an independent sample (N = 49), was .61. Further, a correlation of .59 (N = 163) between the total points at the end of the course and the overall grade-point average provided adequate evidence for considering the classroom tests as satisfactory measures of academic performance. In subsequent analysis Ss were divided into high and low achievement groups depending on whether their total score on achievement tests was above or below the median for 140 Ss.

The 17-item RST data (N = 212) were factor analyzed by means of the principal axis method and normal varimax rotation. Five factors accounting for 66.4 percent of the total variance were retained. Factor 1 (accounting for 16.04 percent of the variance) had high positive loadings on such items as feeling between instructor and students (.81), sense of humor (.77), tolerance and liberality (.74), and ability to arouse interest in students (.48). Factor 2 (accounting for 15.53 percent of the variance) loaded on such items as personal appearance (.81), assignments (.78), personal peculiarities (.61), self-confidence (.56), and scholarship (.50). Factor 3 (accounting for 9.08 percent of the variance) consisted of items such as thinking demanded of students (.85) and leading discussion and questions (.59). Factor 4 had high loadings on items like enunciation (.84) and ability to express thought (.64) and accounted for 10.43 percent of the variance. The last factor had high negative loadings on items such as preparation for class meetings (-.62), interest and enthusiasm in his subject (-.75), organization of the course (-71),

scholarship (-.55), and sense of proportion (-.56). This factor accounted for 15.32 percent of the variance and was reflected in order to facilitate further analysis and interpretation.

Results and Discussion

As a preliminary step it was essential to determine whether the MARS as well as the RST factors were yielding reliable information for the given sample. The corrected split-half reliability coefficients, as reported in Table 1, range between .71 and .93, with a median of .85 (N = 140) for MARS factor scores and of .83 (N = 212) for the RST factors. Subsequently, two-way multivariate analysis of variance (MANOVA) was applied to the ratings of self, father, mother, teacher (during the first week of classes), and teacher (near the end of the semester) on MARS and the teacher ratings (obtained only once at the end of the course) on RST for 140 Ss. The MANOVA results for H_1 (hypothesis of the homogeneity of dispersion matrices) and H_2 (hypothesis regarding the equality of mean vectors) for the main effects of sex and achievement level (AL) and the interaction, Sex x AL, are embodied in Table 1, while Table 2 presents means and standard deviations of MARS and RST factor scores by sex and AL. Since F ratios for H_1, on four of the five ratings on MARS, are not significant at .05, the interpretation of the corresponding F ratios for H_2 is relatively straightforward. Thus, for ratings of self, father, and mother, differences between the two sexes are significant beyond the .0001 level, but the effects of AL and Sex x AL are clearly negligible. These results are compatible with previous findings (Kohn & Fiedler, 1961; Fiedler & Hoffman, 1962) that females rate the significant others more favorably than males, since in the present study the means for females on the desirable dimensions (e.g., productive persistence) are generally higher

than those of males, and for the undesirable characteristics (e.g., unhappiness) the means for the females are generally lower in magnitude than those for the males. It is interesting to note that the differences between males and females on both of the MARS ratings for the teacher follow the same pattern as those for self, father, and mother ratings. However, for teacher ratings, unlike the ratings of self, father, and mother, Ss' achievement level exercises a statistically significant ($p < .05$) influence. The means of the high AL Ss, especially on the dimension of productive persistence, are much higher than those for the low AL Ss. On other MARS factors the differences between high and low AL Ss are generally negligible. The interaction, Sex x AL, is not significant at .05 for any of the ratings of self, father, and mother or for either of the two teacher ratings.

Insert Table 1 about here

Insert Table 2 about here

The RST data provide some additional information regarding the characteristics of Ss vis-a-vis their view of the teacher's effectiveness. First, unlike the MARS data, the F ratio for H_1 is significant at the .01 level, indicating that the variance-covariance matrices for the various groups are not homogeneous. Inspection of the dispersion matrices embodied in Table 3 indicates that the diversity is rather complex. The variance-covariance matrices for males and females are not much different from each other; but for the subgroups, such as females of high AL vs. females of low AL or males of high AL vs. males of low AL, the dispersion differences are considerable. Thus,

females of low AL are appreciably more diverse than females of high AL, while males of low AL are even more diverse than males of high AL on a majority of the RST factors. The low AL males seem to be the most diverse group of all with respect to the variances and covariances of the five RST factors. The significance of the F ratio for H_1, although interesting in itself, should be taken into account in interpreting the F ratios for H_2. However, interpretation of the various F ratios under H_2 is not seriously encumbered by the rejection of H_1 since H_2 is generally insensitive to moderate departures from homogeneity and since rejection of H_1 usually results in a type II error which, in the event of the significance of H_2, becomes a trivial matter.

Insert Table 3 about here

The F ratio for sex differences under H_2 is highly significant ($p < .0001$). This finding corroborates the corresponding result based on the MARS data but contradicts the report that "the sex of student raters bears little or no relationship to their ratings of teachers (Remmers, 1963, p. 368)." Since a high score on any of the five RST factors (including factor 5 which was reflected to conform to the other four in directionality) means that the students regard the teacher favorably, examination of the means indicates that female Ss rated the teacher more favorably than males on all five factors. Similarly, the significant F ratio ($p < .05$) for the differences between mean vectors of high and low AL Ss demonstrates the fact that high AL Ss have a higher regard for the teacher than the low AL Ss. However, the significant interaction ($p < .05$) between sex and AL, when examined in the light of subgroup means, indicates that females with high AL hold the teacher in even higher esteem

than the high AL males while the low AL males rate the teacher even lower than the low AL females. The foregoing findings, although limited by the fact that only one teacher was rated, clearly indicate that both the sex and the achievement level of Ss are among the important determinants of the perceived effectiveness of a college teacher by his students.

References

Fiedler, F. E., & Hoffman, E. L. Age, Sex, and religious background as determinants of interpersonal perception among Dutch children: a cross-cultural validation. Acta Psychologica, 1962, 20, 185-195.

Getzels, J. W., & Jackson, P. W. The teacher's personality and characteristics. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 506-582.

Kerlinger, F. N. Attitudes toward education and perceptions of teacher characteristics. American Educational Research Journal, 1966, 3, 159-168.

Kohn, A. R., & Fiedler, F. E. Age and sex differences in the perception of persons. Sociometry, 1961, 24, 157-164.

McKeachie, W. J. Research on teaching at the college and university level. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 1118-1172.

Quereshi, M. Y. The development of the Michill Adjective Rating Scale (MARS). Journal of Clinical Psychology, 1970, 26, 192-196.

Remmers, H. H. Rating methods in research on teaching. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 329-378.

Yonge, G. D., & Sassenrath, J. M. Student personality correlates of teacher ratings. Journal of Educational Psychology, 1968, 59, 44-52.

Table 1

Results of the Two-way Multivariate Analysis of Variance Based on the MARS (Four Factors) and RST (Five Factors)

MARS Data

Person Rated	F ratios for H_1	F ratios for H_2			Range of Reliability of Factor Scores
		Sex	AL	Sex x AL	
Self	1.23	11.41***	1.59	2.10	.79 to .90
Father	1.34	6.80***	2.14	0.53	.83 to .92
Mother	1.11	11.59***	1.70	1.30	.77 to .92
Teacher (First)	1.51*	11.44***	3.10*	0.94	.79 to .93
Teacher (Last)	0.95	20.72***	5.02**	0.11	.71 to .92

RST Data

Teacher	1.65**	29.73***	2.49*	2.73*	.73 to .85

Note.--For the MARS data df for H_1 and H_2 are 30/50081 and 4/133, respectively; for the RST data, df are 45/45110 for H_1 and 5/132 for H_2.

*$p < .05$. **$p < .01$. ***$p < .0001$.

Table 2

Means and Standard Deviations of MARS and RST Factor Scores by Sex and AL

MARS Factors

Rating	Rater	1		2		3		4		5	
		M	SD	M	SD	M	SD	M	SD	M	SD
Self	M	32.0	7.4	37.7	5.5	38.3	6.2	45.3	6.5		
	F	34.0	9.6	37.4	6.2	34.5	6.9	44.2	7.0		
Father	M	31.7	8.7	35.2	6.0	40.1	7.0	47.5	7.6		
	F	26.6	7.9	35.1	7.6	37.2	8.7	49.3	7.4		
Mother	M	34.9	10.0	38.8	7.5	36.3	8.3	42.7	7.8		
	F	30.7	9.0	38.6	6.5	34.2	8.0	46.3	7.2		
Teacher (First)	M	28.2	8.4	30.2	5.8	38.3	6.4	49.1	7.5		
	F	24.8	6.4	31.2	5.2	38.5	6.6	50.2	6.5		
High AL		28.0	6.9	29.9	5.8	39.0	6.2	49.7	7.4		
Low AL		24.8	8.1	31.6	5.4	37.8	7.3	49.7	6.9		
Teacher (Last)	M	28.4	7.1	30.2	5.3	37.9	5.4	48.7	4.9		
	F	24.8	6.7	31.1	5.1	37.6	5.0	50.2	5.4		
High AL		28.0	6.6	31.2	4.7	37.9	5.1	50.3	4.7		
Low AL		25.1	7.5	30.0	5.7	37.6	5.3	48.6	5.7		

RST Factors

Teacher	M	24.7	4.4	30.4	3.5	14.2	3.0	13.8	3.3	29.9	3.6
	F	25.4	4.8	31.9	2.9	14.5	3.1	14.9	3.2	31.3	3.4
High AL		25.8	5.3	31.4	2.8	14.6	3.2	14.6	3.1	31.3	3.4
Low AL		24.3	4.9	30.9	3.6	14.1	3.0	13.9	3.5	29.8	3.6
High AL Males		25.6	4.3	30.5	2.7	14.7	2.9	14.1	3.0	30.8	3.2
Low AL Males		23.9	4.5	30.3	4.1	13.7	3.2	13.6	3.7	29.0	3.9
High AL Males		26.0	5.2	32.2	2.9	14.5	3.4	15.5	3.1	31.9	3.5
Low AL Females		24.7	4.7	31.7	3.0	14.5	2.9	14.2	3.3	30.6	3.2

Table 3

Variance-Covariance Matrices Based on RST Data

by Sex and AL

Males (Above Diagonal) and Females (Below Diagonal)

Factors	1	2	3	4	5
1	19.77 (23.60)*	5.02	7.32	9.51	9.52
2	10.70	11.79 (8.49)	5.42	2.76	7.95
3	8.60	4.19	9.39 (9.73)	6.43	7.47
4	11.17	5.49	5.24	11.26 (10.61)	6.34
5	12.23	7.02	5.56	7.03	13.05 (11.35)

*Values in parentheses are variances for females.

Males of High AL (Above Diagonal) and of Low AL (Below Diagonal)

1	18.50 (20.15)**	4.97	7.04	8.86	9.40
2	5.02	7.30 (16.47)	1.72	1.00	5.23
3	6.93	9.05	8.62 (9.90)	5.15	5.90
4	10.00	4.49	7.61	9.14 (13.53)	3.53
5	8.43	10.63	8.31	8.80	9.98 (14.87)

**Values in parentheses are variances for low AL males.

Females of High AL (Above Diagonal) and of Low AL (Below Diagonal)

1	26.92 (19.90)***	12.63	10.32	12.12	13.90
2	8.67	8.12 (8.98)	4.90	6.12	7.61
3	7.10	3.53	11.54 (8.15)	4.80	6.63
4	9.62	4.67	5.86	9.82 (10.90)	7.13
5	9.91	6.27	4.63	6.25	12.14 (9.91)

***Values in parentheses are variances for low AL females.

Note.--There are 33 males in high AL and 34 in the low AL group, while there are 37 females in the high AL and 36 in the low AL category.